

高解像度人工衛星画像とAIを用いた 河川域植生図作成手法の開発

DEVELOPMENT OF RIVER AREA VEGETATION MAP PRODUCTION METHOD USING HIGH RESOLUTION SATELLITE IMAGERY AND AI TECHNOLOGIES

宮脇 成生¹・伊川 耕太²・鈴木 研二³・鈴置 由紀洋⁴・池内 幸司⁵
Shigenari MIYAWAKI, Kouta IKAWA, Kenji SUZUKI, Yukihiro SUZUOKI and Koji IKEUCHI

¹非会員 博士(農学) 株式会社建設環境研究所 (〒170-0013 東京都豊島区東池袋2-23-2)

²非会員 修士(林学) 同上

³非会員 博士(農学) 日本スペースイメージング株式会社 (〒104-0031 東京都中央区京橋2-2-1)

⁴非会員 修士(工学・文学) 同上

⁵正会員 博士(工学) 東京大学大学院工学系研究科教授 (〒113-8656 東京都文京区本郷7-3-1)

In this study, we developed a vegetation classification method based on satellite remote sensing, topographical information, and machine learning to reduce the cost and improve the accuracy of vegetation mapping. First, satellite images were processed by object-based classification. Then spectral information (8 spectral bands) from the satellite imagery, the vegetation index, and topographical information were added to each object. Vegetation classification models were created by one of the AI technologies, machine learning algorithms (Random Forests, Support Vector Machine), and vegetation types were classified. As a result, it was confirmed that the concordance rate between the AI classification and the field survey data was higher with 8-band spectral data, and with topographical information.

Key Words : AI, high resolution satellite imagery, remote sensing, riparian topography, riparian vegetation mapping

1. はじめに

河道の樹林化対策, 台風・豪雨に伴う出水による河川の植生変化の把握等, 河川管理における基礎情報として植生図の重要性は高い。この植生図は, 製作者が航空写真から植生の境界線を判別し, さらに現地調査に基づき植生を分類する手法により¹⁾, 多大な労力をかけ作成されている。近年, 衛星リモートセンシング (RS) による植生分類技術の進展もあり²⁾, 植生図の精度向上およびコストダウンの可能性も出てきている³⁾。例えば, 衛星RSの研究分野において, 森林を対象とした植生図作成では, 高解像度 (1m 未満) の衛星画像を用いたオブジェクトベース分類手法や, 機械学習法が導入されている³⁾⁴⁾。しかし, 河川水辺の国勢調査で作成されている植生図 (河川環境基図) には, このような技術は導入されていない。特に, 河川域は草本群落において, 森林を主な領域とする植生図 (例えば, 環境省・自然環境保全基礎調査植生調査⁵⁾) と比較して, 縮尺が詳細であり植

生区分の数も多いため, 衛星RSの導入には検討が必要であろう。また, 近年は衛星画像としてマルチバンドデータ (4バンドあるいは8バンド) の画像入手も可能になっている。これらのデータを土地被覆分類に利用した研究も行われている⁶⁾。

河川域の植生は, 森林域と比較して草地在り占める範囲が広く, 流水による攪乱 (冠水) の影響を受けやすい。攪乱の受けやすさは, 地形 (比高, 水際からの距離等) によって左右される⁷⁾。河川域の植生分類においても, 衛星画像の情報に加えて, この地形情報を考慮することで, 分類精度の向上が期待できる。

そこで, 本研究では, 河川における植生図作成において, 近年の進展が著しい衛星RSおよび地形情報, AI技術の一つである機械学習 (Random Forests, Support Vector Machine) を用いた植生図作成手法を開発した。また, この新規手法で要した調査時間を, 従来手法で要した調査時間と比較した。

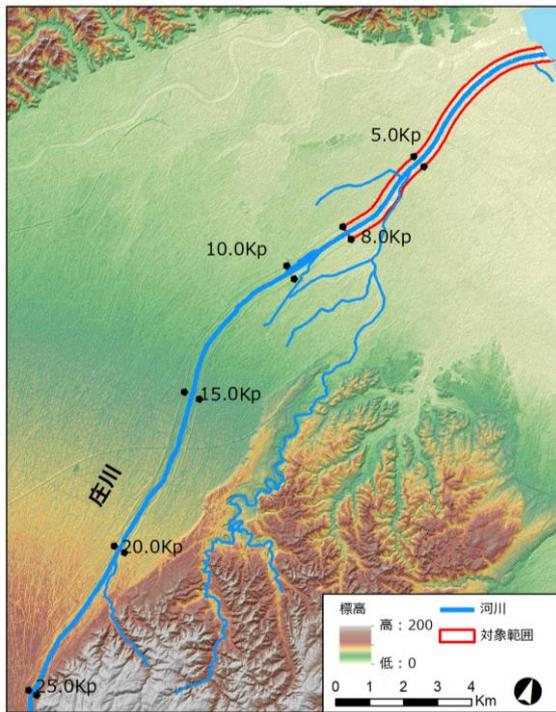


図-1 対象範囲 (背景図データ出典：国土地理院 <http://fgd.gsi.go.jp/download/>).

2. 方法

(1) 対象範囲

対象範囲は、富山県に位置する庄川の河口～8kmの河川区域である(図-1)。ただし、開放水面および人による利用範囲(工事実施箇所、畑・水田などの耕作地、グラウンド・公園等、人工構造物)は、対象範囲から除外した。

(2) 使用データ

解析のフローを図-2に示した。使用したデータについて、以下に示す。

a) 衛星画像のオブジェクト化

オブジェクト化とは、ラスタ画像を比較的均一なピクセルの集合領域である「オブジェクト」に分割することである。本解析に用いた衛星画像は、WorldView-2(地上解像度50cm)により2018年8月14日に取得した。この衛星画像の8バンド(Coastal blue(波長400-450nm), Blue(450-510nm), Green(510-580nm), Yellow(585-625nm), Red(630-690nm), Red edge(705-745nm), NIR1(Near InfraRed 1, 770-895nm), NIR2(Near InfraRed 2, 860-1040nm))のDN(デジタルナンバー)値から、eCognition(Trimble社)を使用してオブジェクト化を行った。形成された各オブジェクトにおいて8バンドのDN値の平均値および3つの植生指数(NDVI(Normalized Difference Vegetation Index 正規化植生指数), NDRE(Normalized Difference RedEdge 正規化レッドエッジ指数), GNDVI(Green

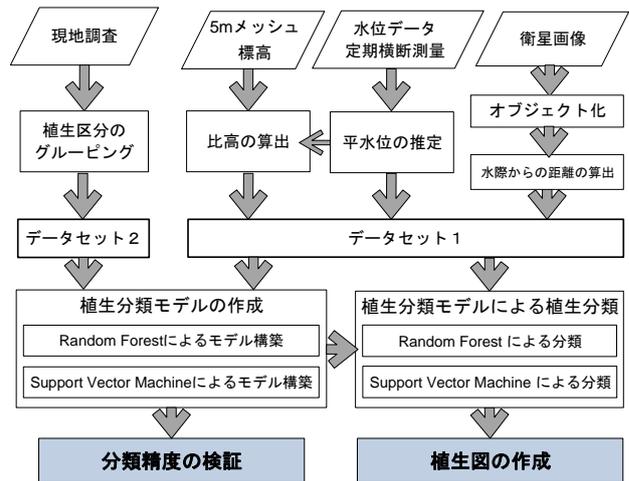


図-2 解析のフロー。

Normalized Difference Vegetation Index 緑正規化植生指数)の平均値を付与した。これら3つの植生指数の算出式(1)～(3)を次に示した。

$$NDVI = \frac{NIR1 - Red}{NIR1 + Red} \quad (1)$$

$$NDRE = \frac{NIR1 - Red\ edge}{NIR1 + Red\ edge} \quad (2)$$

$$GNDVI = \frac{NIR1 - Green}{NIR1 + Green} \quad (3)$$

b) 地形情報

地形情報として、河川の平水位、比高、水際からの距離を算出した。

河川の各距離標における平水位を、宮脇ほか⁹⁾の簡易な推定方法により、各観測所の平水位(2017年)と定期横断測量結果(2018年)を用いて算出した。この各距離標を結ぶ横断線上の平水位を、Natural Neighbor法により河川区域内で補間した。

比高は「地盤高-平水位」と定義した。ここでは、地盤高データとして、基盤地図情報(数値標高モデル)5mメッシュ(標高)(国土地理院ホームページ <http://fgd.gsi.go.jp/download/>よりダウンロード)を用いた。この5mメッシュ標高値と平水位より算出した比高について、各オブジェクト内の平均値を算出し、そのオブジェクトに付与した。

水際からの距離は、本川の開放水面から各オブジェクトまでの最短距離と定義した。この値は、開放水面(本川)と各オブジェクトの最短距離を算出し、各オブジェクトに付与した。

この地形情報と衛星画像に関する情報を付与したオブジェクトデータを「データセット1」(n=6,340)とした。

c) 現地調査

植生分類モデルの学習データとなる植生区分データを取得するため対象範囲において、現地調査を2019年11月

表-1 本研究で用いた植生区分1~4.

植生区分1	植生区分2	植生区分3	植生区分4		
草本	双子葉草本	ヤハズソウ	ヤハズソウ		
		オオブタクサ	オオブタクサ		
		カワラヨモギ・カワラハハコ	カワラヨモギ-カワラハハコ		
		ヨモギ・メドハギ	ヨモギ-メドハギ		
		セイタカアワダチソウ	セイタカアワダチソウ		
		クズ	クズ		
		ノイバラ	ノイバラ		
	単子葉草本	カゼクサ・オオバコ	カゼクサ-オオバコ		
		ヨシ・オギ	カナムグラ	ヨシ	
				オギ	
				クサヨシ	
				ツルヨシ	
		メリケンカルカヤ	メリケンカルカヤ		
		シナダレスズメガヤ	シナダレスズメガヤ		
		低茎単子葉草本	チガヤ	シバ	
				メヒシバ-エノコログサ	
				オヒシバ-アキメヒシバ	
			高茎単子葉草本	オオイヌタデ-オオクサキビ	ススキ
					セイバンモロコシ
				オニウシノケサ	
		湿地植物	ウキヤガラ-マコモ	カンガレイ	
				ヒメガマ	
			湿地植物		
	木本	ヤナギ類	ヤナギ類	ネコヤナギ	
				タチヤナギ	
				ジャヤナギ-アカメヤナギ	
			シロヤナギ		
			カワヤナギ		
低木		低木	アキグミ		
			イタチハギ		
広葉樹		ハンノキ	ハンノキ		
		ムクノキ・エノキ	ムクノキ-エノキ		
		ハリエンジュ	ハリエンジュ		
		広葉樹	ヌルデ-アカメガシワ	ヤマグワ	
				オニグルミ	
針葉樹	クロマツ	クロマツ植林			
タケ類	タケ類	メダケ			
自然裸地	自然裸地	自然裸地	自然裸地		

12~14日に実施した。調査は、1,465地点で実施し、各地点を優占する植物種を記録した。調査時の記録の際、タブレットにはiPad mini (Apple社)、情報入力アプリケーションにはCollectorおよびSurvey123 (ESRI社)を使用した。各地点の位置記録の精度を高めるため、GPSレシーバーGLO2 (Garmin社)をタブレットにBluetooth接続して使用した。

また、この手法により要した現地における調査時間を記録した。この結果を、従来手法(河川水辺の国勢調査)による現地調査の調査時間と比較した。

d) 植生区分のグルーピング

各地点の優占種は、植生区分1~4にグルーピングした(表-1)。最も詳細な(分類クラスの多い)植生区分は、「植生区分4」で43タイプである。他の3つ(植生区分1~3)はこの群落区分4をグルーピングして区分数を少なくしたものを用いた。区分数は、植生区分1が3タイプ、植生区分2が8タイプ、植生区分3が24タイプである。この現地調査で取得した優占種に基づく植生区分を、

表-2 植生分類モデル作成に用いた変数.

分類	変数	植生分類モデル			
		地形情報あり		地形情報なし	
		4バンド	8バンド	4バンド	8バンド
スペクトル密度	Coastal blue		✓		✓
	Blue	✓	✓	✓	✓
	Green	✓	✓	✓	✓
	Yellow		✓		✓
	Red	✓	✓	✓	✓
	Red edge		✓		✓
	NIR1	✓	✓	✓	✓
	NIR2		✓		✓
植生指数	NDVI	✓	✓	✓	✓
	NDRE		✓		✓
	GNDVI		✓		✓
地形情報	比高	✓	✓		
	平水位	✓	✓		
	水面からの距離	✓	✓		

「データセット2」(n=1,465)とした。

(3) 植生分類モデルの作成

データセット1および2を用いて、機械学習の一手法であるRandom Forests(RF)、Support Vector Machine(SVM)により植生分類を行った。

RFは、学習データから複数の決定木を作成し、決定木の多数決で分類モデルを決定する手法である⁹⁾。RFによる植生分類には、R version 3.6.1¹⁰⁾およびパッケージ rangerを用いた。このとき、RFの決定木作成数は500に、作成する決定木の最大階層(max depth)は、10、12、15の3ケースを設定した(以下、各モデルをRF10、RF12、RF15と呼ぶ)。

SVMは、学習データに対して、分類クラス間のマージンが最大となるような超平面を構成することで、判別を行う分類手法である¹¹⁾。SVMによる植生分類には、R version 3.6.1¹⁰⁾およびパッケージe1071を用いた。このとき、予備解析に基づき、カーネル関数には放射基底関数カーネルを用い、Cパラメータは20に設定した。

学習データとしては、データセット1のうちデータセット2の情報がある1,465レコード(現地調査実施地点)のみ解析に用いた。このとき、4つの植生区分のうち1つずつを植生分類モデル作成に用いた。また、各植生分類モデルにおいて、スペクトル密度、植生指数、地形情報は、表-2の組み合わせで用いた。

さらに、植生分類モデル作成に用いるデータ数を1,465、1,000、750、500、250に変化させ、モデルによる分類結果と現地調査結果の一致率を比較した。すべてのデータを用いる場合(データ数1,465)以外では、①データが10未満の植生区分については全データをサンプリングし、残りの植生区分についてはデータをランダムにサンプリング、②植生分類モデルを作成、③植生分類モデルによる分類結果と現地調査の一致率を算出、この①~③を10回繰り返して、一致率の平均値を算出した。

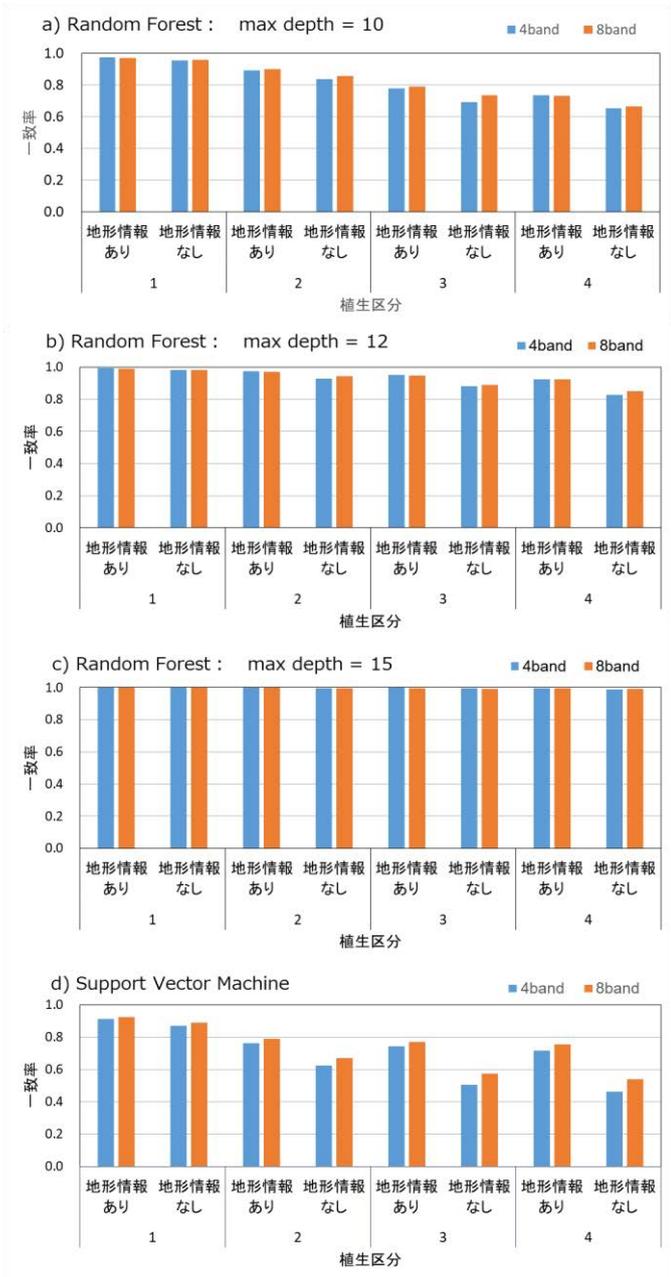


図-3 各植生分類モデルによる分類と現地調査との一致率。

(4) モデルによる植生分類

作成した植生分類モデルとデータセット1を用いて、対象範囲の植生区分を推定し、その結果に基づき植生図を作成した。

3. 結果

データセット1および2 (n=1,465) を使用し、機械学習の手法別 (RF10, RF12, RF15, SVM) に、地形情報の有無、衛星画像のバンド数 (4バンド, 8バンド) で一致率の比較を行った (図-3)。機械学習手法では、RF15, RF12の一致率が高く、RF10, SVMの一致率が低かった。植生区分については、いずれの手法でも区分数が少ない

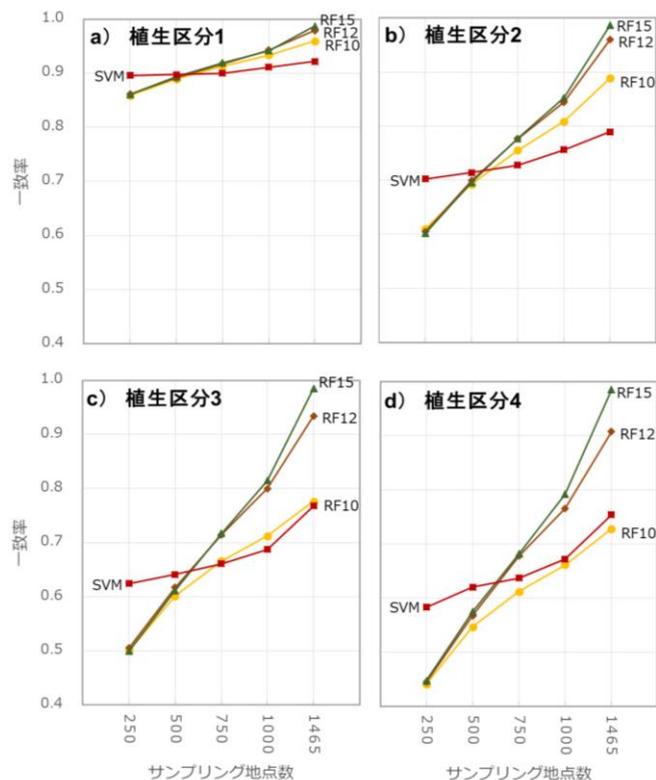


図-4 植生分類モデル (8バンド, 地形情報あり) 構築に用いたサンプリング地点数と一致率 (平均値) の関係。

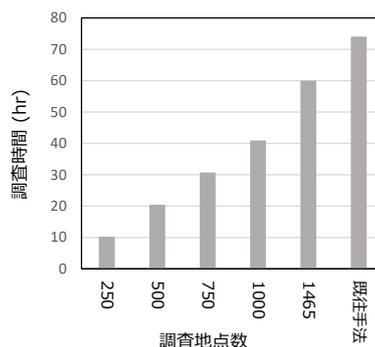


図-5 調査地点数と調査時間の設定条件。

ほど、一致率は高かった。地形情報については、いずれの手法でも、地形情報ありのデータの一致率が高かった。衛星画像のバンド数についても、いずれの手法でも、4バンドと比較して8バンドのデータの一致率が高かった。

植生分類モデルを作成する際のデータ数と一致率の関係については、衛星画像は8バンドおよび地形情報ありのデータで、手法はRF10, RF12, RF15, SVMを用いた場合、いずれのモデルにおいても、サンプリング地点数の減少とともに一致率は低下した (図-4)。いずれのモデル、サンプリング地点数でも、植生区分1から4に向けて一致率は低下した。しかし、植生区分3と4の差は小さかった。また、いずれの植生区分でも、RF15, RF12, RF10の順で一致率は低下した。SVMモデルは、サンプリング地点数が500, 250の場合、いずれの植生区分においてもRFより一致率が高かった。

本手法の現地調査に要した時間は60時間で、従来手法

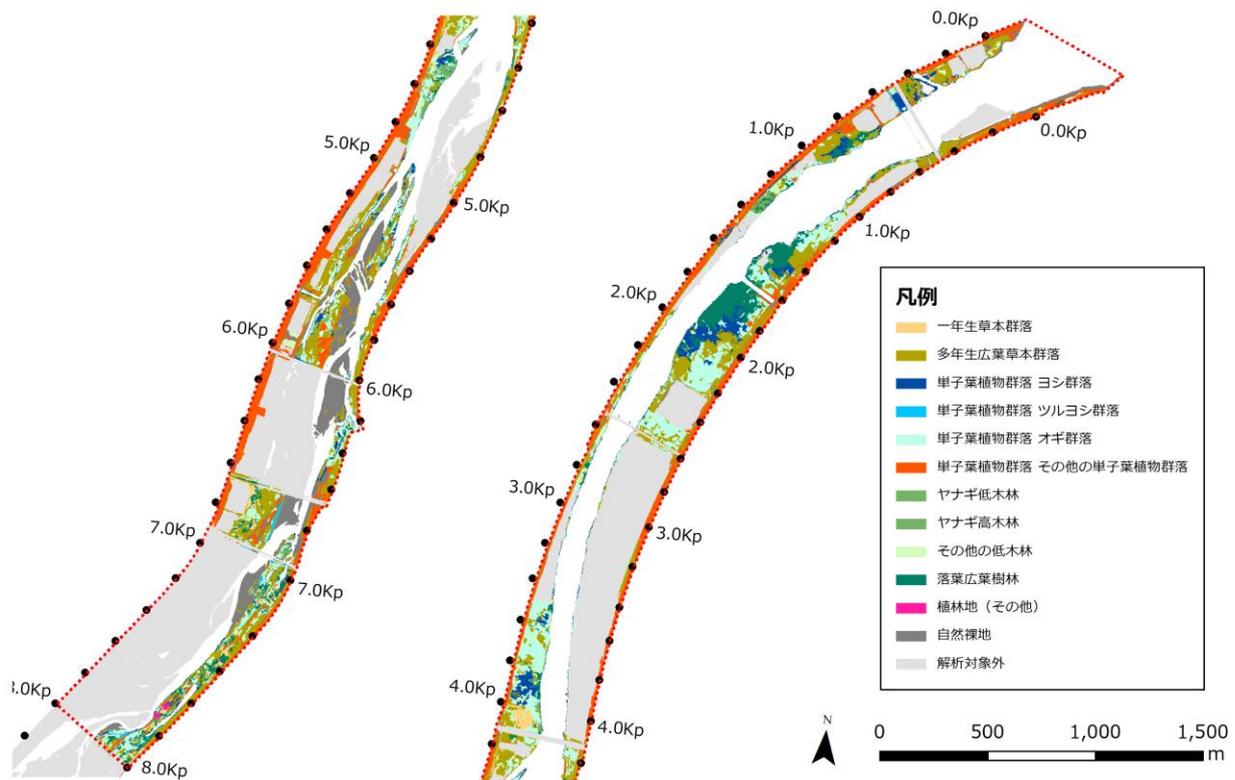


図-6 提案手法により作成した庄川の植生図の例 (RF15モデル, 植生区分4) .

は74時間と比較して19%少なかった (図-5) . 本調査は、タブレットおよび入力アプリケーションの操作に不慣れた調査員が実施したこと、調査地点の設定に戸惑ったことにより、1地点あたりの調査時間を過剰に費やしてしまった。

作成した植生分類モデル64モデルのうち、植生区分4のRF15による8バンド・地形情報ありデータを用いたモデルの植生図を示した (図-6) .

4. 考察

本研究では、河川域における植生図作成に、衛星リモートセンシング技術を導入した手法の開発を行った。具体的には、高解像度衛星画像 (地上解像度0.5m) と地形データ (比高等) と植生の対応関係を機械学習 (RF, SVM) により植生分類モデル化し、このモデルから対象範囲の植生分類を行う手法を提案した。その結果、衛星画像のスペクトルデータは、4バンドと8バンドデータを比較すると8バンドデータを使用した方が、また地形情報を使用した方が、現地調査データとの一致率が高くなることが確認された。また、今回の現地調査に要した時間は、従来手法と比較して約20%減であり、より簡易な分類の植生区分ではさらに少ない時間で調査が可能であることが示された (図-4, 図-5) .

本研究で使用した調査地点の地形情報 (平水位, 比高, 水際からの距離) は、河川における縦横断方向における位置および増水による冠水のしやすさの指標となってい

る。河川における既往の研究では、比高に関わる指標と植物の分布の相関が示されている⁷⁾。本研究において地形情報を用いた植生分類モデルがより高い一致率を示したことは、これを裏付けるものであると考えられる。河川域の植生は、山地の森林域と異なり、頻繁に生じる洪水による攪乱の影響を受けるため、リモートセンシングによる植生分類を行う際にも、本研究で用いた地形情報は、分類精度を高める上で重要な役割を果たすと考えられる。

既往の植生図作成手法で要する労力が大きい現地調査にかかるコストを削減することができれば、植生図作成全体のコストが削減されることが期待できる。現地調査に関わるコスト削減を考える際、植生分類モデルの学習データ数を変化させたときの一致率との関係が参考になる。学習データ数が小さくなるということは、現地調査にかかる労力が小さくなることを表す。現地調査の労力を減らしたときに一致率は変動する。この一致率の許容ラインを設定することで、現地調査の労力を設定できると考えられる。この点について、既往研究での十分な検討が行われていないため、本研究で解析を行った。RFを用いた場合、最大階層が10 (RF10) と15 (RF15) のモデルを比較すると、学習データ数が多い場合 (データ数1,465) はRF15のモデルが非常に高い一致率を示すが、学習データ数を減らすとモデル間で一致率の差は小さくなった。

RFとSVMの一致率を比較すると、学習データ数が多いモデルでは圧倒的にRF15が高い一致率を示したが、学習データ数が少ない場合はSVMが高い一致率を示し

た。ただし、対象とする河川植生の状況（全データ数、植生区分毎のデータ割合等）により結果は異なる可能性があるため、他河川での検討がさらに必要であろう。

リモートセンシングによる植生分類において、モデルの学習および正確性の評価には、十分な数のデータ数が必要であると考えられる。このデータ数は、使用できる時間、予算、労働力により制約される。しかし、必要なデータ数についての経験則は存在せず、多くの場合データ数が多ければ多いほどよいと考えられている¹²⁾。より多くのデータを用いれば、モデルを複雑にすることができる。その一方で、植生区分の分類が複雑な場合（例えば、河川域）ではより多くのデータが必要だろう¹²⁾。

本研究では、リモートセンシング手法を用いた植生図作成手法を提案した。現地調査については、今後この調査方法に習熟することで、1地点あたりの作業時間の削減が期待できるだろう。また、調査地点数の設定により、要する調査時間の大幅な削減が期待できる。例えば、調査地点数1,000地点で作成するならば、従来調査から調査時間を46%削減できる。一方、この手法は従来手法と用いる技術が異なることに加えて、作成の工程も大きく異なる（図-7）。従来手法で利用する航空写真の撮影は、多くの場合、現地調査実施の前年度となるだろう。提案手法では、衛星画像の取得を現地調査と同じ年としている。もし、現地調査実施の年度に写真を撮影すれば、現地調査における植生境界の修正は、その大部分が不要となるだろう。従来手法における植生境界の現地調査における判別・修正も正確性から問題がある。例えば、平面的な位置の正確性、現地における植生区分に関わる人間認識のバイアスの問題が挙げられる。一方、本研究における提案手法を用いるならば、これらの問題は解決あるいは部分的に軽減される。また、衛星画像をオブジェクト化する作業において、データに対して同一パラメータを設定すれば、再現性のある植生境界線を作成できる。

一方、本研究において沈水植物群落は、衛星画像からの抽出を行っていないため、今後の検討が必要である。

以上のように、本研究では河川域における植生図作成において、本手法の有効性を示すことができたと考えられる。また、これらの現地調査におけるコスト減に加えて、植生図作成に関わる室内作業についても、従来手法では手作業だった部分も、提案手法では作業量を大幅に削減できると考えられる。今後はさらに大規模な河川における試行を含め、現地データを使用しての検証を積み重ねていくことが必要であろう。

謝辞：本研究で用いた定期横断測量、河川水辺の国勢調査の現地調査時間は、国土交通省北陸地方整備局富山河川国道事務所より提供いただいた。また、清水元子さんには、GISによるデータ処理を手伝っていただいた。ここに記して御礼申し上げる。

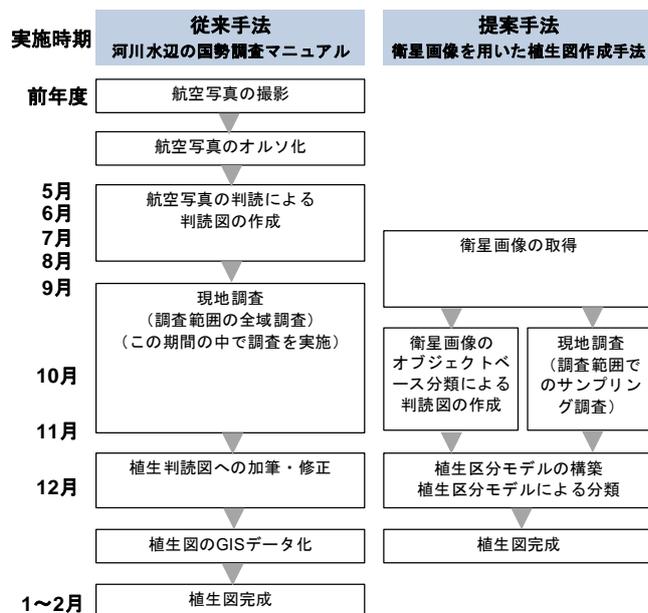


図-7 衛星画像を用いた植生図作成手法の実実施スケジュール。

参考文献

- 1) 国土交通省水管理・国土保全局河川環境課: 平成28年度版河川水辺の国勢調査基本調査マニュアル[河川版] (河川環境基図作成調査編), 2016.
- 2) 村上拓彦, 望月翔太: リモートセンシングによる植生マッピング, 日本生態学会誌, 64, 233-242, 2014.
- 3) 鎌形哲稔, 原慶太郎: オブジェクトベース画像解析による植生図化における分類精度の評価, 植生学会誌, 27, 83-94, 2010.
- 4) 櫻井聖悟, 伊藤達夫, 田中和博: 高分解能衛星データを用いたオブジェクトベース林相区分における地形情報の利用, 日本林学会誌, 91, 1-8, 2009.
- 5) 環境省自然環境局生物多様性センター: 自然環境保全基礎調査植生調査実施要領 (平成29年度版), 2017.
- 6) 橋本秀太郎・田殿武雄・小野里雅彦・堀雅裕・塩見慶: 多時期光学観測データを用いた高精度土地被覆分類手法の開発, 日本リモートセンシング学会誌, 34, 102-112, 2014.
- 7) 宮脇成生, 鷲谷いづみ: 河川域における侵略的外来植物4種の侵入範囲予測, 保全生態学研究, 15, 17-28, 2010.
- 8) 宮脇成生, 永山滋也, 加藤康充, 伊藤英恵, 萱場祐一: 観測所水位データから平常時水位の縦断形を推定する〜庄内川を事例として〜, 応用生態工学, 21, 53-60, 2018.
- 9) Breiman, L.: Random Forests, *Machine Learning*, 45, 5-32, 2001.
- 10) R Core Team R: *A language and environment for statistical computing*, R Foundation for Statistical Computing, 2018.
- 11) Vapnik, V.: *Statistical Learning Theory*, Wiley, 1998.
- 12) Wohlfart, C., Bevanda, M., Horning, N., Leutner, B. and Wegmann, M.: Field data for remote sensing data analysis, *Remote Sensing and GIS for Ecologists: Using Open Source Software*, Wegmann, M., Leutner, B. and Dech, S. eds., Pelagic Publishing, pp.136-149, 2016.

(2020. 4. 2受付)